

**PARTIAL CREDIT MODEL (PCM) DALAM
PENSKORAN POLITOMI PADA
TEORI RESPON BUTIR**

SKRIPSI



Oleh :
SAFARUDDIN
H 121 07 006

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2013**

**PARTIAL CREDIT MODEL (PCM) DALAM PENSKORAN
POLITOMI PADA TEORI RESPON BUTIR**

Skripsi

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana

Sains pada Jurusan Matematika Fakultas Matematika dan Ilmu

Pengetahuan Alam Universitas Hasanuddin, Makassar



Oleh :
SAFARUDDIN
H 121 07 006

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2013**

LEMBAR KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan
sesungguh-sungguhnya bahwa skripsi yang saya buat dengan

judul:

**“PARTIAL CREDIT MODEL (PCM) DALAM
PENSKORAN POLITOMI PADA
TEORI RESPON BUTIR”**

adalah benar hasil kerja saya sendiri, bukan hasil plagiat dan belum
pernah dipublikasikan dalam bentuk apapun.

Makassar, 26 Maret 2013

SAFARUDDIN
NIM : H 121 07 006

**PARTIAL CREDIT MODEL (PCM) DALAM
PENSKORAN POLITOMI PADA TEORI
RESPON BUTIR**

Disetujui oleh

Pembimbing Utama

Pembimbing Pertama

Anisa, S.Si, M.Si
NIP: 19730227 199802 2 001

Drs. M. Saleh AF, M.Si
NIP: 19540804 197802 1 001

Pada Tanggal: 26 Maret 2013

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN**

Pada hari ini, tanggal 26 Maret 2013, Panitia Ujian Skripsi menerima dengan baik skripsi yang berjudul :

**“PARTIAL CREDIT MODEL (PCM) DALAM PENSKORAN
POLITOMI PADA TEORI RESPON BUTIR”**

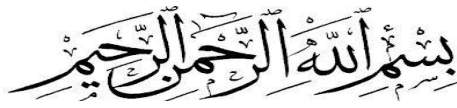
yang diajukan untuk memenuhi salah satu syarat guna memperoleh gelar Sarjana Sains pada Program Studi Statistika Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Makassar, 26 Maret 2013

Panitia Ujian Sripsi

		Tanda Tangan
1. Ketua	Drs. H. Muhammad Hasbi, M.Sc	()
2. Sekretaris	Hendra, S.Si, M.Kom	()
3. Anggota	Dra. Nasrah Sirajang, M.Si	()
4. Anggota	Anisa, S.Si, M.Si	()
5. Anggota	Drs. M. Saleh AF, M.Si	()

KATA PENGANTAR



Alhamdulillah Rabbil Alamin puji syukur penulis panjatkan ke hadirat **Allah SWT** atas limpahan rahmat dan ridho-Nya yang tak terhingga serta kesempatan dan kesehatan yang dikaruniakan-Nya kepada penulis sehingga skripsi ini dapat terselesaikan dengan baik. Shalawat dan salam tak lupa penulis haturkan kepada baginda Rasulullah **Muhammad SAW** manusia terbaik serta para sahabat dan pengikut jejak Beliau yang selalu menjadi panutan segenap umat manusia, semoga kebaikan dan keselamatan juga terus tercurah hingga akhir zaman.

Skripsi ini merupakan salah satu persyaratan dalam menyelesaikan pendidikan di Program Studi Statistika Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin, Makassar. Dalam penyelesaian skripsi ini, penulis telah melewati perjuangan panjang dan pengorbanan yang tidak sedikit. Namun berkat rahmat dan izin **Allah SWT** serta dukungan dan bantuan dari berbagai pihak baik secara moril maupun materil, langsung atau tidak langsung, sehingga akhirnya skripsi ini dapat terselesaikan. Untuk itu penulis sangat tulus dan bangga menyampaikan ucapan terima kasih dan penghargaan yang tak terhingga kepada Ayahanda **La Ebo. U, S.Pd** dan Ibunda **Wa Beda** yang telah membesarkan dan mendidik penulis dengan penuh cinta dan kasih sayang yang sangat tulus serta dengan penuh kesabaran dan kesungguhan dalam memberikan motivasi dan dukungan moril dan materil serta doa yang tak henti-hentinya demi keberhasilan penulis selama menjalani proses pendidikan. Untuk saudara-saudaraku **Sertu Darsalam, Bripda Sawaluddin, Al Akbar** dan **Nudia Fazira** yang selalu memberikan dorongan dan motivasi tersendiri kepada penulis juga terima kasih atas dukungan semangat dan doa yang telah kalian berikan.

Penghargaan yang tulus dan penuh keikhlasan penulis ucapkan terimakasih kepada :

- **Ibu Anisa, S.Si, M.Si** selaku pembimbing utama dan **Bapak Drs. M. Saleh AF, M.Si** selaku pembimbing pertama atas kesediaan dan kesabaran untuk membimbing dan membagi ilmunya serta nasehat yang tak ternilai harganya kepada penulis dalam penyusunan skripsi ini.
- **Ibu Dr. Hasmawati, M.Si**, selaku Ketua Jurusan Matematika, atas ilmu dan nasehatnya. **Bapak Dr. Nurdin, M.Si** selaku Sekretaris Jurusan Matematika yang juga telah memberikan banyak bantuan dan dukungan. **Ibu Dra. Nasrah Sirajang, M.Si** selaku Penasehat Akademik penulis yang selalu memberikan dorongan, nasehat, dan bantuan yang tak ternilai harganya. Para **Dosen Jurusan Matematika** yang telah mencurahkan ilmunya dan mendidik penulis selama di bangku kuliah hingga berhasil menyelesaikan studi.
- Para **Staf Jurusan Matematika (Pak Sutamin, S.Sos dan Pak Nasir)**, atas bantuan, pelayanan dan kerjasamanya yang diberikan kepada penulis.
- Teman-teman **Statistika 2007** (Agil, Wandu, Fadli, DJ, Juhu, Sutha, Ilo, Tika, Lilies, Azlam, Ranu, Midel, Nandang, Iqbal, Ekky, Pingky, Dian, Echy, Uchy, Dilla, Lila, Kmute, Ovi, Alm. Marwa, Ayyub, Nengnonk, Nila, Mega, Bunda, Ada'). Dan teman-teman **Ekspansi'07**. Terima kasih atas dukungan dan kebersamaan kalian kawan.
- Teman-teman pondokan (La Kasih, La Oval, La Mangkana, La Faris, La La Haris, La Pandu, La Tuna, La Gamsir, La Ogo, La Tinus, K'Arif, La Suda, La Asraf, La Risal, La Birun, La Bretek, La Hery, La Gengky, La Budi, La Hendrik, La Ical, La Hendro, La Jamil, La Node, La Iman, La Sukri, dan La Abas) yang selalu menemani hari-hariku selama menyelesaikan studi. Thank's bro.
- Kanda-kanda senior, adik-adik dan seluruh warga HIMATIKA yang tidak dapat penulis sebut satu per satu.
- Kanda-kanda senior, teman-teman, dan adik-adik di komisariat IPPERMATO-MAKASSAR yang juga tidak dapat penulis sebutkan satu per satu.

- Semua pihak yang telah membantu melancarkan proses penyelesaian skripsi ini yang tidak dapat penulis sebutkan satu per satu

Semoga segala bantuan, dukungan dan partisipasinya dapat bernilai ibadah dan mendapat pahala yang setimpal. Akhir kata, semoga tulisan ini dapat memberikan manfaat bagi semua pihak yang membutuhkan dan terutama bagi penulis. *Amin Yaa Rabbil Alammin.*

Makassar, 28 Februari 2013

Penulis

ABSTRAK

Pada skripsi ini kasus yang dibahas adalah penskoran politomi dengan teori respon butir pada soal ujian Mid Semester Mata Kuliah Matematika Dasar yang akan diberikan kepada mahasiswa Fakultas Kehutanan Universitas Hasanuddin tahun ajaran 2011/2012 dengan menggunakan metode partial credit model, juga memuat model matematika, estimasi parameter, dan estimasi kemampuan peserta tes. Dalam pelaksanaan tes uraian, penskoran biasanya dilakukan secara parsial berdasarkan langkah-langkah yang harus ditempuh untuk menjawab benar suatu butir soal. Penskoran dilakukan perlangkah dan skor perbutir diperoleh peserta dengan menjumlah skor siswa tiap langkah, dan kemampuan diestimasi dengan skor mentah. Model penskoran seperti ini belum tentu tepat, karena tingkat kesulitan tiap langkah tidak diperhitungkan. Pendekatan alternatif yang dapat digunakan yaitu pendekatan teori respon butir (TRB) untuk penskoran politomi, salah satunya dengan *partial credit model* (PCM). PCM merupakan pengembangan dari model Rasch pada butir dikotomi yang berisi satu parameter lokasi butir dan dengan PCM kemudian dikembangkan dengan menjabarkan lokasi butir menjadi kategori. Skor kategori pada PCM menunjukkan banyaknya langkah untuk menyelesaikan dengan benar butir soal tersebut, sehingga kemampuan tiap peserta tes dapat diestimasi dengan menghitung probabilitas tiap peserta dalam menjawab tiap langkah dalam menyelesaikan sebuah soal tes. Dan hasil dari penelitian ini didapat bahwa butir soal yang layak dipakai untuk uji tes pada mahasiswa Fakultas Perikanan Unhas tahun ajaran 2011/2012 adalah butir soal nomor 1 dan 7 untuk butir soal X, dan soal nomor 2,7,8,dan 12 untuk butir soal Y.

Kata kunci : teori respon butir, *partial credit model*

ABSTRACT

.In this thesis the cases discussed are scoring polytomous item response theory in the Mid Semester exam Basic Mathematics Lecture will be given to students of the Faculty of Forestry, University of Hasanuddin academic year 2011/2012 using the partial credit model, also includes mathematical modeling, parameter estimation , and the estimate of the test participants' ability. In the description of the test execution, scoring is usually done partially based on the steps that must be taken to correctly answer a question item. Scoring done by the pace and score points obtained by adding up scores of participants with students every step, and the ability to estimate the raw scores. Such scoring model is not necessarily appropriate, because the level of difficulty of each step is not taken into account. An alternative approach that can be used that approach to item response theory (IRT) for scoring polytomous, one with partial credit model (PCM). PCM is a development of Rasch models in point dichotomy that contains the item location parameter and the PCM then developed to describe the location of item categories. Score category at the PCM shows the number of steps to resolve the matter of the items correctly, so the ability of each test taker can be estimated by calculating the probability of each participant to answer each step in solving a test problem. And the results of this research obtained that the grains matter of who worthy worn for test tests on students of Faculty of Fisheries University of Hasanuddin academic year 2011/2012 is the grain problem number 1 and 7 for the grain problem X, and matter of numbers 2,7,8, and 12 for grains matter of Y.

Keywords: item response theory, *partial credit models*

DAFTAR ISI

HALAMAN JUDUL	i
KATA PENGANTAR	ii
ABSTRAK	iv
ABSTRACT	v
DAFTAR ISI	vi
DAFTAR GAMBAR	viii
DAFTAR TABEL	ix
DAFTAR LAMPIRAN	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA	6
2.1 Teori Respon Butir (<i>Item Response Theory</i>)	6
2.2 Karakteristik Tes dalam Teori Respon Butir (TRB).....	8
2.3 Rancangan Tes pada Teori Respon Butir.....	10
2.4 Model Politomi <i>Partial Credit Model (PCM)</i>	11
2.5 Uji Kesamaan Variansi.....	15
BAB III METODOLOGI PENELITIAN	18
3.1 Sumber Data	18

	3.2 Variabel Penelitian	18
	3.3 Metode Analisis Data	18
BAB IV	HASIL DAN PEMBAHASAN	20
	4.1 Rancangan Penarikan Sampel	20
	4.2 Mengestimasi Parameter butir	20
	4.3 Uji Levene.....	28
	4.4 Menghitung Nilai Peluang dengan PCM	29
BAB V	PENUTUP	36
	5.1 Kesimpulan	36
	5.2 Saran	36
	DAFTAR PUSTAKA	41

DAFTAR GAMBAR

Gambar 4.1 Diagram Rekapitulasi nilai tiap kategori perangkat tes X.....	23
Gambar 4.2 Diagram Rekapitulasi nilai tiap kategori perangkat tes Y.....	24
Gambar 4.3 Grafik Tabel nilai tingkat kesukaran tahap (δ_{ij}) dan tingkat kesukaran butir(b) perangkat X.....	26
Gambar 4.4 Grafik Tabel nilai tingkat kesukaran tahap (δ_{ij}) dan tingkat kesukaran butir(b) perangkat Y.....	27
Gambar 4.5 Grafik Nilai P peserta K1 pada soal X.....	33
Gambar 4.6 Grafik Nilai P peserta K2 pada soal Y.....	34

DAFTAR TABEL

Table 4.1 Tabel Skor Peserta perangkat tes X.....	21
Table 4.2 Tabel Skor Peserta perangkat tes Y	22
Table 4.3 Rekapitulasi nilai tiap kategori perangkat tes X.....	23
Table 4.4 Rekapitulasi nilai tiap kategori perangkat tes Y.	23
Table 4.5 Tabel nilai tingkat kesukaran tahap (δ_{ij}) dan tingkat kesukaran butir(b) perangkat X.....	25
Table 4.6 Tabel nilai tingkat kesukaran tahap (δ_{ij}) dan tingkat kesukaran butir (b) perangkat Y.....	26
Tabel 4.7 Klasifikasi tingkat kesukaran butir perangkat tes X dan Y.....	27
Tabel 4.8 Tabel 4.8 Tabel nilai signifikan uji levene.....	28
Tabel 4.9 Nilai P peserta K1 pada soal X.....	31
Tabel 4.10 Nilai P peserta K2 pada soal Y.....	33

DAFTAR LAMPIRAN

Lampiran 1. Table Data Mentah Score Hasil Ujian Mahasiswa.....	38
Lampiran 2. Hasil Output SPSS Uji Levene.....	40

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Tes merupakan salah satu cara paling mudah dan murah yang bisa dilakukan untuk memotret kemajuan belajar peserta tes dalam ranah kognitif. Oleh karena itu, keberadaan perangkat tes yang berkualitas merupakan suatu keniscayaan sehingga kemampuan kognitif peserta tes dapat diungkapkan (Djunaidi, 2008). Dalam bidang pendidikan, kegiatan penilaian atau evaluasi hasil belajar peserta didik merupakan salah satu tugas penting yang harus dilakukan oleh pendidik. Evaluasi hasil belajar peserta didik dilakukan untuk mengetahui kemajuan peserta didik terhadap kurikulum yang telah diajarkan. Namun terkadang perangkat tes yang digunakan dalam mengukur kemampuan Responden yang berbeda tidak mencerminkan prestasi belajar yang sebenarnya. Banyak metode yang dapat digunakan untuk menganalisis perangkat tes mulai dari pendekatan klasik hingga modern. Pendekatan secara klasik yang digunakan adalah teori tes klasik (*classical true-score theory*) sedangkan pendekatan modern dengan teori respon butir (*item response theory*) (Vitaria, 2010).

Teori tes klasik telah lama menjadi acuan dalam pembuatan alat ukur atau instrumen di dunia pendidikan. Dalam teori ini setiap *item* ditelaah menurut tingkat kesukaran *item* dan daya beda *item*. Namun, karakteristik *item* tersebut tergantung pada kelompok sampel mana yang akan digunakan atau dikenai tes. Begitu juga pada koefisien reliabilitas dan validitas tes akan menjadi lebih tinggi

apabila kelompok subyek merupakan kelompok yang kemampuannya heterogen (bervariasi besar) dan tes yang sama akan mempunyai koefisien yang lebih rendah apabila dikenakan pada kelompok yang kemampuannya relatif homogen. Ketergantungan pada kelompok subyek ini tentu akan mengurangi manfaat parameter *item* dalam berbagai aplikasinya. Kelemahan lain, dalam teori tes klasik diperlukan asumsi kesetaraan eror pengukuran bagi semua subyek yang dikenai tes. Keberatan asumsi ini adalah kurangnya dukungan yang memperkuatnya dikarenakan pada tes yang sulit, eror pengukuran bagi subyek yang berkemampuan rendah akan berbeda dari eror bagi subyek yang berkemampuan tinggi (Azwar, 2007).

Keterbatasan dan kelemahan tersebut menjadi dasar dalam pengembangan teori tes baru dalam dunia pengukuran modern yang dapat melengkapi dan memperbaiki teori tes klasik, yang disebut sebagai teori respon butir (*item response theory*, IRT) atau teori sifat laten (*latent trait theory*). Teori respon butir merupakan teori pengukuran modern yang digunakan dalam menganalisis *item*. Teori ini mempunyai orientasi pada *item* yang karakteristiknya tidak tergantung pada kelompok tertentu. Teori respon *item* membebaskan ketergantungan antara *item* tes dan peserta tes (konsep invariansi parameter), respon peserta tes pada satu *item* tes tidak mempengaruhi *item* tes lainnya (konsep independensi lokal), dan *item* tes hanya mengukur satu dimensi ukur (konsep unidimensional). Sehingga aplikasinya menjawab kebutuhan dunia pengukuran modern hingga saat ini, yaitu perbandingan antar kemampuan peserta, pengembangan bank *item*, bahkan pengembangan tes adaptif, dan lain-lain (Vitaria, 2010).

Bentuk soal dalam tes prestasi belajar, secara umum dapat dikelompokkan menjadi dua kategori yaitu: 1) tes uraian, terdiri dari uraian bebas, uraian terbatas atau isian singkat, uraian berstruktur, dan 2) tes objektif, terdiri dari pilihan benar salah, pilihan ganda, dan menjodohkan. Kualitas tes, termasuk bentuk tes pilihan ganda (dikotomi) dapat diungkap melalui analisis butir soal secara teoretis (telaah) dan analisis empiris. Analisis butir soal secara kualitatif dilakukan untuk menilai butir soal ditinjau dari aspek materi, konstruksi, dan bahasa. Analisis secara kuantitatif menekankan pada analisis karakteristik butir soal secara empiris. Karakteristik butir soal antara lain meliputi indeks kesukaran (p), daya beda (d), dan distribusi respons.

Dalam pelaksanaan tes uraian, penskoran dilakukan perlangkah dan skor perbutir diperoleh peserta dengan menjumlah skor siswa tiap langkah, dan kemampuan diestimasi dengan skor mentah. Model penskoran seperti ini belum tentu tepat, karena tingkat kesulitan tiap langkah tidak diperhitungkan. Selain itu, peluang menjawab benar seorang siswa berdasarkan respons tertentu tidak dapat diprediksikan. Pendekatan alternatif yang dapat digunakan dengan pendekatan teori respons butir untuk penskoran politomi. Penskoran politomi adalah pemberian skor pada hasil tes yang terdiri dari dua nilai atau lebih, dimana penskorannya dilakukan langkah perlangkah dalam suatu butir soal, sehingga dalam proses analisisnya memperhitungkan tingkat kesulitan pada tiap langkah dalam menyelesaikan butir soal tersebut. Skor tertinggi tentu saja didapatkan ketika peserta tes mampu menyelesaikan dengan benar soal hingga langkah akhir. Ada beberapa model yang bisa digunakan dalam analisis butir politomi, salah

satunya dengan *partial credit model* (PCM). Berdasarkan uraian di atas, penulis mencoba mengkaji tentang “ ***Partial Credit Model (PCM) dalam Penskoran Politomi pada Teori Respon Butir***”

1.2. Rumusan masalah

Dalam penulisan tugas akhir ini, permasalahan yang dibahas yaitu bagaimana melakukan penskoran butir dengan metode *Partial Credit Model(PCM)*, mengestimasi parameter dan mengestimasi kemampuan peserta tes.

1.3. Batasan masalah

Berdasarkan permasalahan di atas, maka ruang lingkup penelitian ini dibatasi pada penskoran politomi pada butir dalam bentuk soal urian dengan metode *Partial Credit Model.(PCM)*, dan dengan menggunakan rancangan tes ekuivalen.

1.4. Tujuan penelitian

1. Mengestimasi parameter butir, indeks kesukaran butir dan daya beda butir.
2. Melakukan penskoran politomi pada butir dengan metode *Partial Credit Model*
3. Menentukan butir soal yang layak untuk peserta tes.

1.5. Manfaat penelitian

Penelitian tugas akhir ini dapat bermanfaat sebagai sumbangan informasi, pemikiran mengenai penerapan ilmu statistika khususnya penggunaan teori respon butir dengan penskoran politomi dalam peningkatan mutu pendidikan serta diharapkan juga dapat bermanfaat bagi aspek akademis.

BAB II

TINJAUAN PUSTAKA

2.1. Teori Respons Butir (*Item Response Theory/IRT*)

Teori Respon Butir adalah teori yang menyatakan tentang suatu pembuatan perangkat ukur yang analisisnya berdasarkan bentuk lengkungan kurva karakteristik butir yang diperoleh dari hasil respon para peserta tes. Dengan menggunakan teori respon butir kemampuan peserta tes dapat dievaluasi dan seberapa baik kemampuan suatu butir soal dalam suatu tes dapat dideskripsikan. Teori respon butir digunakan untuk mengatasi kelemahan yang terjadi pada teori tes klasik. Perhitungan pada teori tes klasik mengaitkan tingkat kesukaran dengan kemampuan kelompok peserta tes secara langsung, sedangkan analisis pada teori respon butir tingkat kesukaran butir tidak dikaitkan langsung dengan kemampuan peserta tes tetapi tingkat kesukaran butir dikaitkan langsung dengan karakteristik butir. Analisis butir tes dilakukan dengan tujuan mengetahui atau mencari butir tes yang berkualitas untuk digunakan sebagai alat pengukur baik untuk tes hasil belajar maupun penelitian yang berkaitan dengan kognitif. Menurut Hambleton, Swaminathan, & Rogers (1991: 5) secara umum ciri-ciri teori respons butir adalah sebagai berikut:

1. Karakteristik butir tidak tergantung pada peserta ujian,
2. Skor yang digambarkan peserta ujian tidak tergantung pada tes,
3. Merupakan model yang lebih menekankan pada tingkat butir dari pada tingkat tes,

4. Merupakan model yang tidak mensyaratkan secara ketat tes paralel untuk menaksir reliabilitas, dan
5. Merupakan model yang menguraikan sebuah ukuran keputusan untuk tiap skor kemampuan yakni ada hubungan fungsional antara peserta tes dengan tingkat kemampuan yang dimiliki.

Selain itu, menurut Hambleton, Swaminathan, & Rogers (1991: 9) mengemukakan bahwa asumsi-asumsi yang melandasi teori respons butir adalah:

1. Unidimensi, setiap butir hanya mengukur satu ciri peserta.
2. Independensi lokal, respon pada butir (*item*) yang satu bebas dari pengaruh respon pada butir lain jika kemampuan yang mempengaruhi performansi dibuat konstan.
3. Fungsi karakteristik butir atau kurva karakteristik butir, merefleksikan hubungan yang sebenarnya antara kemampuan dan respon peserta terhadap butir tes.

Ditinjau dari banyaknya asumsi, teori respons butir memang diharapkan akan memberikan informasi yang lebih teliti dibandingkan teori tes klasik. Namun demikian, berkaitan dengan ukuran sampel, baik teori tes klasik maupun teori respons butir tidak mensyaratkan ukuran sampel secara pasti.

Parameter-parameter dalam Teori respon Butir (*Item response Theory*) antara lain,

1. tingkat kesulitan *item* (b),
2. daya beda *item* (a),
3. peluang tebakan semu (c),

4. parameter peserta (θ),
5. respon peserta terhadap *item* dinyatakan dalam bentuk probabilitas menjawab benar ($P_i(\theta)$).

2.2. Karakteristik Tes dalam Teori Respon Butir (TRB)

Karakteristik butir soal ditunjukkan oleh parameter kemampuan yang digunakan dalam model teori respon itu. Untuk mengetahui sejauh mana keterterapan teori respon, maka perlu diadakan pengujian terhadap parameter kemampuan yang digunakan dalam teori tersebut. Secara umum karakteristik butir test yang umumnya digunakan adalah untuk menentukan apakah butir soal itu layak atau tidak digunakan untuk mengukur kemampuan peserta test (Responden). Karakteristik butir test biasanya digunakan dua bentuk parameter yaitu sebagai berikut:

2.2.1. Tingkat kesukaran butir soal (b)

Tingkat kesukaran soal adalah kemampuan peserta ujian secara keseluruhan untuk menjawab butir soal dengan benar. Ada beberapa alasan untuk menyatakan tingkat kesukaran soal. Bisa saja tingkat kesukaran soal ditentukan oleh kedalaman soal, kompleksitas, atau hal-hal lain yang berkaitan dengan kemampuan yang diukur oleh soal. Namun demikian, ketika kita mengkaji lebih mendalam terhadap tingkat kesukaran soal, akan sulit menentukan mengapa sebuah soal lebih sukar dibandingkan dengan soal yang lain. Intinya, bermutu atau tidaknya butir-butir *item* tes hasil belajar pertama-tama dapat diketahui dari derajat kesukaran atau taraf kesukaran yang dimiliki oleh masing-masing butir

item tersebut. Butir-butir *item* tes hasil belajar dapat dinyatakan sebagai butir-butir *item* yang baik, apabila butir-butir *item* tersebut tidak terlalu sukar dan tidak pula terlalu mudah dengan kata lain derajat kesukaran *item* itu adalah sedang atau cukup. Angka yang dapat memberikan petunjuk mengenai tingkat kesulitan *item* itu dikenal dengan istilah *difficulty index* (angka indeks kesukaran *item*), yang dalam dunia evaluasi hasil belajar umumnya dilambangkan dengan huruf *b*.

Tingkat kesukaran butir soal dapat diperoleh dengan cara berikut:

$$b_i = \frac{n_i}{N_i} \quad (2.1)$$

b_i = tingkat kesukaran butir *i*

n_i = jumlah skor yang diperoleh peserta yang menjawab pada butir *i*

N_i = skor maksimal pada butir *i*

Namun pada bentuk butir politomi, nilai tingkat kesukaran butir adalah jumlah dari tingkat kesukaran tahap tiap butir.

Sedangkan untuk menentukan tingkat kesukaran tahap pada suatu item dapat diperoleh dengan cara berikut:

$$\delta_{ij} = \frac{n_{ij}}{N_i} \quad (2.2)$$

δ_{ij} = tingkat kesukaran tahap *j* pada butir *i*

n_{ij} = jumlah skor yang diperoleh peserta yang menjawab pada butir *i* kategori *j*

N_i = skor maksimal pada butir *i*

Kategori tingkat kesukaran soal yaitu sebagai berikut :

- a. Tingkat kesukaran antara 0,00 – 0,30 soal tergolong sulit artinya hanya 0 – 30% siswa yang menjawab benar.

- b. Tingkat kesukaran antara 0,31 – 0,70 soal tergolong sedang artinya hanya 31 – 70% siswa yang menjawab benar.
- c. Tingkat kesukaran antara 0,71 – 1,00 soal tergolong mudah artinya 71 – 100% siswa yang menjawab benar, (Adolf, 2009).

2.3. Rancangan Tes pada Teori Respon Butir

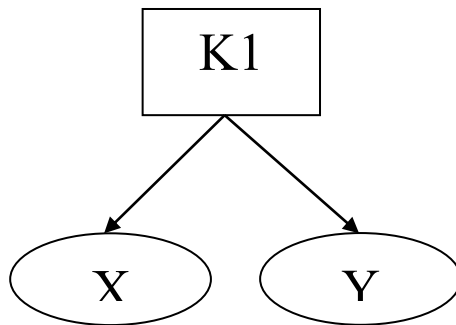
Petersen dkk (dalam Holland & Rubin, 1982) mengemukakan tiga jenis rancangan dalam penyetaraan skor diantara beberapa tes yang berbeda, yaitu Rancangan Kelompok Tunggal (A), Rancangan Kelompok Ekuivalen (B), dan Rancangan Tes Jangkar (C). Berikut adalah skema rancangan tersebut.

Tabel 2. 1. Skema Rancangan tes

Jenis rancangan	Kelompok peserta tes	
	K ₁	K ₂
A	X	Y
B	X ,Y	
C	X +R	Y + R

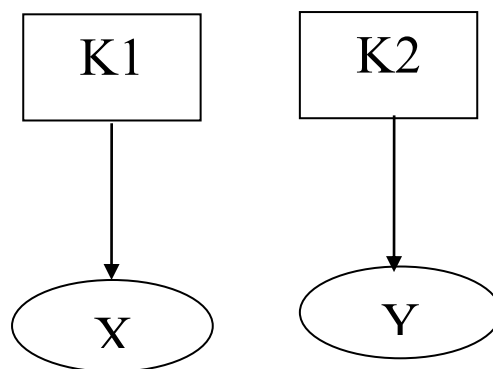
Keterangan : diadaptasi dari Crocker & Algina, 1986.

Pada rancangan kelompok tunggal, kegiatan penyetaraan dilakukan dengan menggunakan satu kelompok peserta yang merespons dua perangkat tes misalnya X dan Y. Parameter butir dari kedua perangkat tes diestimasi secara terpisah dengan mengkalibrasi parameter kemampuan peserta atau parameter butir. Berdasarkan rancangan ini, dengan mengkalibrasi parameter kemampuan peserta maka parameter butir dari perangkat tes X dan Y sudah berada pada skala yang sama (Hidayati, 2002). Untuk ilustrasi:



Gambar 2.1 Rancangan Kelompok Tunggal

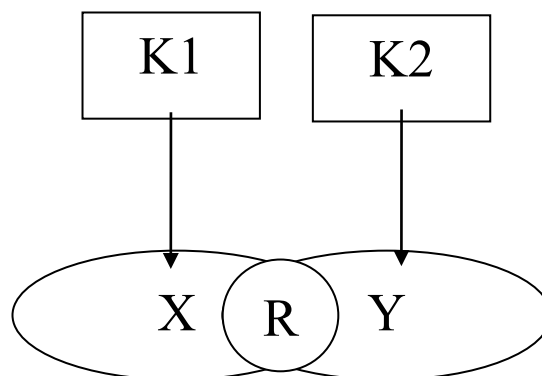
Pada rancangan Ekuivalen, dua perangkat tes diberikan pada dua kelompok peserta yang sama kemampuannya atau ekuivalen. Proses secara spiral digunakan dalam desain ini, dimana peserta tes dibagi dua secara acak kemudian masing-masing mendapat perangkat tes X dan tes Y. Sebagai ilustrasi, misalnya terdapat dua kelompok K1 dan K2 dan dua perangkat tes misalnya X dan Y. kelompok K1 mengerjakan perangkat tes X dan kelompok K2 mengerjakan perangkat tes Y.



Gambar 2.2 Rancangan Ekuivalen

Rancangan tes jangkar, biasanya didesain jika masalah keamanan tes

menjadi salah satu pertimbangan penting dan memungkinkan untuk menyelenggarakan beberapa tes dalam satu waktu. Pada desain ini masing-masing perangkat tes mempunyai beberapa item yang sama (common item) dan masing-masing kelompok mengerjakan perangkat tes yang berbeda. Kelompok peserta tes tidak dibentuk atas dasar acak. Pada rancangan ini, peserta tes hanya disyaratkan ada dua kelompok, yaitu K_1 dan K_2 . Selanjutnya, perangkat tes X dan Y masing-masing ditambah dengan perangkat tes pengait atau perangkat tes jangkar atau anchor (R), sehingga perangkat tes menjadi $X + R$ dan $Y + R$. Setelah itu, K_1 mengerjakan perangkat tes $X + R$ dan K_2 mengerjakan $Y + R$ yang akhirnya ditemukan formula konversi.



Gambar 2.3 Rancangan Anchor

Namun pada penelitian ini kami hanya menggunakan rancangan B atau rancangan kelompok ekuivalen (*Equivalent-Group Design*).

2.4 Model Politomi *Partial Credit Model (PCM)*

Menurut Mardapi (1991:7), pada awalnya teori respons butir menggunakan distribusi normal, namun dalam perkembangan selanjutnya

digunakan model distribusi logistik. Hal ini dikarenakan model distribusi logistik lebih sederhana analisis matematikanya. Teori respon butir memiliki dua ciri utama, pertama, teori respon butir memiliki variabel respon berupa data dikotomi yaitu 1 (untuk respon dari peserta tes yang benar) dan 0 (untuk respon dari peserta tes yang salah). Kedua, teori respon butir memiliki kurva karakteristik butir berbentuk seperti huruf S. Kedua ciri tersebut senada dengan model logistik, khususnya model logistik biner. Model logistik biner merupakan model yang digunakan untuk menganalisis hubungan antara satu variabel respon dan beberapa variabel bebas, dengan variabel responnya berupa data kualitatif dikotomi. Kurva fungsi distribusi model logistik menyerupai huruf S. Kesamaan ciri tersebut menjadi dasar pengembangan model pada teori respon butir menggunakan model logistik. Dalam teori respon butir digunakan model matematis untuk menghubungkan karakteristik butir dengan kemampuan peserta tes. Hubungan tersebut digambarkan melalui kurva karakteristik butir.

Terdapat beberapa model pengukuran yang termasuk kategori teori respon butir. Model pengukuran tersebut dibedakan berdasarkan jumlah parameter butir yang dimasukan kedalam model yaitu model logistik satu parameter, model logistik dua parameter, dan model logistik tiga parameter. Perbedaan ketiga model tersebut terletak pada banyaknya parameter yang digunakan untuk menggambarkan karakteristik butir dalam model yang bersangkutan. (Hambleton dkk, 1991: 7)

Model logistik memiliki satu parameter ciri peserta dan satu sampai tiga parameter ciri butir yang tergabung ke dalam bentuk fungsi yang menentukan

probabilitas jawaban benar dalam suatu butir tertentu. Namun untuk dapat mengetahui nilai pada model logistik ini, perlu diketahui berapa besar nilai parameter itu dalam suatu pengukuran. penentuan nilai parameter, ini dikenal sebagai mengestimasi parameter. Jika yang diperhatikan peserta pada suatu pelaksanaan uji tes, maka pengestimasian ini dilakukan untuk mengetahui bagaimana ciri peserta atau bagaimana kemampuan peserta pada uji tes itu. Pengestimasian ini bersangkutan dengan ciri peserta atau dengan kemampuan peserta. Tapi jika yang diperhatikan butir dalam ujian tes maka pengestimasian ini dilakukan untuk mengetahui bagaimana ciri butir itu. Pengestimasian ini bersangkutan dengan ciri butir atau taraf kesukaran butir, perbedaan butir dan kebetulan menjawab dengan benar pada butir itu.

PCM merupakan pengembangan dari model Rasch butir dikotomi yang diterapkan pada butir politomi. Model Rasch butir dikotomi yang berisi satu parameter lokasi butir kemudian dikembangkan dengan menjabarkan lokasi butir menjadi kategori. Persamaan Model Rasch menurut Han & Hambleton (2007:15) dituliskan sebagai berikut:

$$P_i(\theta) = \frac{1}{1+e^{-D(\theta-b_i)}} \quad (2.7)$$

Nilai peluang setiap peserta berhasil mengerjakan item i merupakan fungsi logistik perbedaan parameter kemampuan θ dengan tingkat kesukaran item b_i . Persamaan diatas dapat ditulis kembali sebagai berikut:

$$P_i(\theta) = \frac{1}{1+e^{-D(\theta-b_i)}} = \frac{\exp(D(\theta-b_i))}{1+\exp(D(\theta-b_i))} = \frac{P_{i1}}{P_{i0}(\theta)+P_{i1}(\theta)} \quad (2.8)$$

Sehingga persamaan RM untuk peserta n dan item i dengan skor x sebesar 0 dan 1 dengan kemampuan sebesar θ dan tingkat kesukaran item sebesar δ adalah sebagai berikut (Masters, 1999:101;Wright, 1982:39-40).

$$P_{nix} = \frac{1}{1+\exp(\theta_n-\delta_{i1})} \quad \text{untuk } x = 0 \quad (2.9)$$

dan

$$P_{nix} = \frac{\exp(\theta_n-\delta_{i1})}{1+\exp(\theta_n-\delta_{i1})} \quad \text{untuk } x = 1 \quad (2.10)$$

Model diatas merupakan model estimasi pada bentuk soal dengan skala penskoran dikotomi. Kemudian dikembangkan pada skala politomi yang memiliki pola skor x sebesar $0,1,2,3,\dots,m_i$. Sehingga peluang seorang peserta pada tingkat kemampuan θ meraih skor x diatas $x-1$ dapat dihitung dengan persamaan sebagai berikut (Han & Hambleton, 2007:15).

$$\frac{P_{ix}(\theta)}{P_{ix-1}(\theta)+P_{ix}(\theta)} = \frac{\exp(D(\theta-b_{ix}))}{1+\exp(D(\theta-b_{ix}))} \quad \text{untuk } x = 0,1,2,3, \dots, m_i \quad (2.11)$$

$P_{ix}(\theta)$ dan $P_{i(x-1)}(\theta)$ mengacu pada peluang seorang peserta dengan kemampuan θ meraih skor x dan $x-1$. Sehingga peluang seorang peserta dengan kemampuan θ untuk memperoleh skor x pada item i dengan tingkat kesukaran item sebesar δ dapat dituliskan sebagai berikut:

$$P_{nix} = \frac{1}{1+\exp(\theta_n-\delta_{ij})} \quad \text{untuk } x = 0 \quad (2.12)$$

$$P_{nix} = \frac{\exp(\theta_n-\delta_{ij})}{1+\exp(\theta_n-\delta_{ij})} \quad \text{untuk } x = 1,2,3 \dots, m_i \quad (2.13)$$

Keterangan:

$P_{nix}(\theta)$ = peluang peserta ke- n dengan tingkat kemampuan θ memperoleh skor x

pada butir i

θ_n = tingkat kemampuan peserta

δ_{ij} = tingkat kesukaran tahap (j) pada butir (i).

Dengan demikian, tingkat kesukaran butir untuk butir i sebesar δ akan terurai menjadi nilai delta sebesar δ_{ij} untuk $x=1,2,3,\dots,m_i$. Butir nomor satu memiliki tiga kategori atau diskors secara politomi tiga kategori, memiliki δ_{11} dan δ_{12} , butir nomor 2 memiliki δ_{21} δ_{22} . Besarnya nilai delta-1 menunjukkan nilai yang diperlukan peserta untuk berpindah dari kategori 1 (skor 0) ke kategori 2 (skor 1) dan nilai delta-2 menunjukkan nilai yang diperlukan untuk berpindah dari kategori 2 (skor 1) ke kategori 3 (skor 2). Besarnya delta-1 dapat lebih kecil, sama, atau lebih besar dari delta-2.

Parsial Kredit Model (PCM) pada skala politomi yang merupakan pengembangan dari model Rasch pada skala dikotomi. Asumsi pada PCM yakni setiap butir mempunyai daya beda yang sama, namun indeks kesukaran dalam setiap langkah tidak perlu terurut, suatu langkah dapat lebih sulit dibandingkan langkah berikutnya. Skor kategori pada PCM menunjukkan banyaknya langkah untuk menyelesaikan dengan benar butir tersebut. Skor yang lebu tinggi menunjukkan kemampuan yang lebih besar daripada skor kategori yang lebih rendah. Jika i adalah butir politomi dengan kategori skor x sebesar $0,1,2,\dots,m_i$, maka probabilitas dari individu n skor x butir i dapat dituliskan dalam persamaan berikut:

$$P_{nix}(\theta) = \frac{\exp \sum_{j=0}^x (\theta - \delta_{ij})}{\sum_h^{m_i} \exp \sum_{j=0}^h (\theta - \delta_{ij})} \quad \text{untuk } x = 0,1,2, \dots, m_i \quad (2.14)$$

Keterangan:

$P_{nix}(\theta)$ = peluang peserta ke-n dengan tingkat kemampuan θ memperoleh skor x
pada butir i

x = skor peserta

j = kategori/tahap dalam butir soal

i = butir soal

n = peserta

θ_n = tingkat kemampuan peserta ke n

δ_{ij} = tingkat kesukaran tahap (j) pada butir (i).

Persamaan di atas dapat dijabarkan berdasarkan jumlah kategori di dalam butir. Misalnya sebuah butir memiliki kategori dengan skor 0,1,2. Maka kita dapatkan kategori (j) sebanyak 3 buah persamaan yang probabilitas individu pada tiap kategori.

$$\text{Probabilitas pada kategori 0 } P_{i0}(\theta) = \frac{1}{1 + \exp(\theta_n - \delta_{i1}) + \exp[(\theta_n - \delta_{i1}) + (\theta_n - \delta_{i2})]} \quad (2.5)$$

$$\text{Probabilitas pada kategori 1 } P_{i1}(\theta) = \frac{\exp[(\theta_n - \delta_{i1})]}{1 + \exp(\theta_n - \delta_{i1}) + \exp[(\theta_n - \delta_{i1}) + (\theta_n - \delta_{i2})]} \quad (2.6)$$

$$\text{Probabilitas pada kategori 2 } P_{i2}(\theta) = \frac{\exp[(\theta_n - \delta_{i1}) + \exp[(\theta_n - \delta_{i2})]}{1 + \exp(\theta_n - \delta_{i1}) + \exp[(\theta_n - \delta_{i1}) + (\theta_n - \delta_{i2})]} \quad (2.7)$$

Probabilitas pada kategori 0 terlihat ada angka 1 pada bagian penyebut. Hal ini dikarenakan dalam PCM mensyaratkan persamaan berikut

$$\sum_{j=0}^2 (\theta - \delta_{ij}) = 1$$

2.5 Uji Kesamaan Variansi

Uji kesamaan variansi atau uji homogenitas dimaksudkan untuk memperlihatkan bahwa dua atau lebih kelompok data sampel berasal dari populasi yang memiliki variansi yang sama. Kehomogenan varian untuk beberapa kelompok data dapat diperiksa melalui box-plot. Selain dengan melihat box-plot, kehomogenan variansi dapat diuji dengan statistik uji F, uji Bartlett dan uji Levene. Uji F dilakukan jika kelompok sampel tidak lebih dari dua. Untuk data yang lebih dari dua kelompok sampel digunakan uji Bartlett dan uji Levene. Statistik uji Bartlett dapat dipergunakan jika data berdistribusi normal, sedangkan statistik uji Levene tidak membutuhkan syarat kenormalan data (Supra, 2009). Karena dalam penelitian ini hanya menggunakan dua kelompok sampel maka digunakan uji Levene.

Pengujian hipotesis dua variansi dilakukan untuk mengetahui variansi dua populasi sama atau tidak (Afgani, 2009).

Jika S_1^2 dan S_2^2 merupakan penduga σ_1^2 dan σ_2^2 maka rumus varian:

$$S_1^2 = \frac{\sum X_i^2}{k_1-1} - \frac{(\sum X_i)^2}{k_1(k_1-1)} \quad (2.8)$$

$$S_2^2 = \frac{\sum Y_i^2}{k_2-1} - \frac{(\sum Y_i)^2}{k_2(k_2-1)} \quad (2.9)$$

dimana,

S_1^2 = variansi dari parameter peserta kelompok tes pertama (X) dengan k_1 butir soal

S_2^2 = variansi dari parameter peserta kelompok tes kedua (Y) dengan k_2 butir soal.

Hipotesis yang digunakan yaitu,

H_0 : variansi parameter peserta kelompok tes X dan Y sama ($\sigma_1^2 = \sigma_2^2$)

H_1 : variansi parameter peserta kelompok tes X dan Y tidak sama ($\sigma_1^2 \neq \sigma_2^2$)

F_{tabel} ditentukan berdasarkan taraf nyata (α), derajat bebas pembilang ($v_1 = k_1 - 1$), dan derajat bebas penyebut ($v_2 = k_2 - 1$)

Adapun kriteria pengujiannya yaitu,

H_0 diterima, jika $F_{1-\frac{1}{2}\alpha(v_1;v_1)} < F_{hitung} < F_{\frac{1}{2}\alpha(v_1;v_1)}$

H_0 ditolak, jika $F_{hitung} \leq F_{tabel} = F_{1-\frac{1}{2}\alpha(v_1;v_1)}$ atau $F_{hitung} \geq F_{tabel} = F_{\frac{1}{2}\alpha(v_1;v_1)}$

dimana,

$$F_{1-\alpha(v_1;v_1)} = \frac{1}{F_{\alpha(v_1;v_1)}} \text{ maka, } F_{1-\frac{1}{2}\alpha(v_1;v_1)} = \frac{1}{F_{\frac{1}{2}\alpha(v_1;v_1)}}$$

Dimana uji statistik,

$$F_{hitung} = \frac{S_1^2}{S_2^2}$$

dengan derajat bebas pembilang adalah v_1 dan derajat bebas penyebut adalah v_2 .